

SEQUENCE LISTING

(i) GENERAL INFORMATION:

(1) APPLICANT: Cen, Chong Jin  
Lim, Gek Keow  
Leong, Ai Lin  
Zhao, Yi  
Chen, Wei Ning

(ii) TITLE OF INVENTION: A VACCINE INDUCED HEPATITIS B VIRAL  
STRAIN AND USES THEREOF

(iii) NUMBER OF SEQUENCES: 11

(iv) CORRESPONDENCE ADDRESS

(A) ADDRESSEE: Ladas & Parry  
(B) STREET: 26 West 61 Street  
(C) CITY: New York  
(D) STATE: New York  
(E) COUNTRY: USA  
(F) ZIP: 10023

(v) COMPUTER READABLE FORM.

(A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC DOS/MS-DOS  
(D) SOFTWARE: PatentIn Release #1.0, Version #1.33

(vi) CURRENT APPLICATION DATA

(A) APPLICATION NUMBER: 09/719,533  
(B) FILING DATE: 13-DEC-2000  
(C) CLASSIFICATION: 435

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: PCT/SG98/00045  
(B) FILING DATE: 19-JUN-1998

(viii) ATTORNEY AGENT INFORMATION:

(A) NAME: Marc Clifford J.  
(B) REGISTRATION NUMBER: 30,086  
(C) REFERENCE/LOCKET NUMBER: U 013108.9

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (212) 703-1890

(x) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 213 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: circular

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CTCCACCACT TTCCAGCAAA CTETTCAGA TCCCAGAGTC AGGGCCCTGT ACTTTCCTGC 1  
 TGGTGGCTCC AGTTCAGGAA CAGTGAGGCC TGCTCAGAAT ACTGTCTCTG CCATATCTGC 120  
 AATCTTATCG AAGACTGSGG ACCCTGTAGC GAACATGGAG AACATGGCAT CAGGACTCCT 180  
 AGGACCCCTG CTCGTGTTAC AGGCGGGGTT TTTCTTGTG AAAAAATCC TCACAATACC 240  
 GCAGAGTCTA GACTCGTGGT GGACTTCTCT CAATTTTCTA GGGGGAACAG CCGTGTGTCT 300  
 TGACCAAAAT TCGCAGTCCC AAATCTTAG TCACTCAGCA AATGTTCTT CTTAACTTG 360  
 TCTGTGTTAT CGCTGGAGGT GTGTGGGGG TTTATATAT TCTCTCTGCA TCTGTGTCT 420  
 ATGECTCATC TTCTTGTTGG TTCTTCTGGA CTATCAAGGT ATGTTGCCCG TTTGTCTCT 480  
 AATTCAGGA TCAACAACAA CCAGCACCGG ACCATGCAAA ACCTGCACAA CTCTGCTCA 540  
 AGGAACCTCT ATGTTTCCCT CATGTTGCTG TACAAAACCT ACGGACAGAA ACTGCACCTG 600  
 TATTCGATC CCATCATCTT GGGCTTTGG AAAATAGCTA TGGGAGTGGG CTTGAGTCCG 660  
 TTTCTCTTGG CTCAGTTTAC TAGTGCCATT GTTTCAGTGG TTCTAGGGG TTTCCCCAC 720  
 TGTCTGGCTT TCAGTTATAT GGATGATGTT GTTTTGGGG CCAAGCTCTT ACAACATCTT 780  
 GAGGCTCTT ATGCTCTCT TACCACTTT CTCTTATCTT TATGATCA TTTTATCTT 840  
 CATAAAACA AAGATGGGG ATATTCCTT AACTTCATG CATATCTCAT TGGGATTTGG 900  
 GGCACATTC CACAGGAAAT TATTGTACA AAAATCAAAA TGTCTTTAG GAAATTTCTT 960  
 GTAAAGAGGG ATATTGATG GAAATATCT CACCAATTT TGGTCTTTT GGGTTTGGC 1020  
 GCGCTTTCA TCAATGTGG ATATCTCTCT TTAATGCTT TATATGATG TATACAAGCA 1080  
 AAACAGGCTT TACTTTTCT GCAAACTTAC AAGACTTTC TAAATAAACA GATCTGAACT 1140  
 TTTAGCTTA TTCTTGGTA AAGCTTCTT CTCTGCAAA TGTCTCTTA GCAAACTTCT 1200  
 ACTGCTTGG GCTTGGCCAT AGGCATCAAT CGATGCTG GAACTTTCT GTCTCTCTG 1260  
 GCGATCLATA CTGGGGAAT CTATGCGCT TTTTCTCT GCAAGAGCT TGGGSCAAAA 1320  
 CTGATCAGGA CTGACAATT TCTCTGCTG TCGCGAAGT ATACATATT TCAATGCTG 1380  
 CTATGCTGCT CTAAATCT GATCTGCGG GAGAGTCTT TTCTTATCT CCAATCTCT 1440  
 CTCAATCTCG CGGACGAGCT CTCTGGGGG CTTTCTGGG TCTATCTCTG GCTTCTCTCG 1500  
 TTTTATATT TAATCTCT GAGGCGAAT TTTCTTACT CGGACTCTT GTCTCTCTT 1560  
 TCTCATTTG CGGACCGTCT GACTTCTCT TCACTCTG AGTCTGATG GAGACCACTG 1620  
 TGAACGCGCA CGGAAACCT CCAAGGTCT TCAATAAGAG GACTCTTGA CTTTCAGCAA 1680

TGTCAAGGAC CACCTTGAG GCATACTTCA AAGACTGTGT GTTTAATGAG TGGGAGGAGT 1740  
 TGGGGGAGGA GGTAGGTTA AAGGTCTTTG TACTAGGAGG CTGTAGGCAT AAATTGCTGT 1800  
 GTTCACCATC AGCATGGAAC TTTTTCACCT CTGCCTAATC ATCTCATGTT CATGTCTTAC 1860  
 TGTTCGAAGCC TCCAAGCTGT GCCTTGGGTG GCTTTGGGGC ATGGACATTG ACCCGTATAA 1920  
 AGAATTTGGA GTTCTGTGG AGTTACTCTC TTTTTTGCCT TCTGACTTTT TTCTTTCTAT 1980  
 TTTAGATCTC TTTTATGCG CTTCTTTTCT GTATGGGAG GCCTTAGAGT CTCTTTAACA 2040  
 TTCTTACCT CAGATACGG CACTCAGGCA AGCTATTCTC AGTTGGGGTG AGTTAATGAA 2100  
 TCTAGGCATC TGGGTGGGAA GTAATTTGGA AGATCCAGCA TCCAGGGAAT TAGTAGTCAG 2160  
 CTATGTCAAC GTTAATATGG GCCTAAAAAT CAGACAACCTA TTGTGTTTTT ACATTTCTCTG 2220  
 TCTTACTTTT GGGAGAGAAA CTGTTCTTGA ATATTTGGTG TCTTTTGGAG TGTGGATTCTG 2280  
 CACTCTCTCT GCATATAGAC CACCAATGC CCCTATCTTA TCAACACTTC GGGAAACTAC 2340  
 TCTTCTTA GAAGAGGGA GGTCCCTAG AAGAAGAACT CCTCGCCTC GCAGACGAAG 2400  
 GTCTCAATCG CCGCTGCGA GAAGATCTCA ATCTCGGGAA TCTCAATGTT AGTATTCCTT 2460  
 GGCACATAA GGGGGAAAG TTTAGGGGG GTTATTCCTC TACGTTACT TCTTTTATC 2520  
 GTAAATGGA AACTCTTCT TTTCCGACA TTCAATTTGA GGAGGACAT CTGTATAGAT 2580  
 GTAAGCAAT TGGGGGGCC CTACAGTAA ATGAAAACAG GAGACTAAAA TTAATTATGC 2640  
 CTGCTAGTT TTATCAAAAT GTTACTAAAT ATTTGCCCTT AGATAAGGG ATCAAAACAT 2700  
 ATTATCCAGA GTATGTAGTT AATCATTAAT TCCAGACCGG ACATTATTTA CACACTCTTT 2760  
 CGAAGGGCGG GATCTTATAT AAAAGAGAAT CCACACGTAG CGCCTCATT TCGGGCTCAC 2820  
 CATATTCTTG GAATTAACAT CTACAGCATG GGAGTTTGGT CTTCGAAACC TCGAAATGGG 2880  
 ATGGGGAATA ATCTTCTGT CCCCATTCT CTGGATTTCT TCCCGATCA TCAGTTGGAT 2940  
 CCTGCATTCA AAGTCAACTC AGAAAAATCA GATTGGGAGC TCAATGGCA CAAGGACAAC 3000  
 TGGCGGAGG GGAACAAGCT GGGAGTGGCA GCATTGGGG CAGGTTTAC CCGTCTTCA 3060  
 TGGGGACTCT TGGGTGGAG CCTTCAGCT CAGGCTTAC TCACAACCTGT GCGAGCAGCT 3120  
 TCTCTCTCTG CTTTCAGCA TGGCACTCA GGAAGGACG CTACTGCTT ATCTCAGCT 3180  
 TAAATTA TCAATTTA GTTATGAG TGGAA 3240

2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 843 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```

Met  Pr  Leu  Ser  Tyr  Gln  His  Phe  Arg  Lys  Leu  Leu  Leu  Leu  Asp  Glu
1      5      10      15

Glu  Ala  Gly  Pro  Leu  Glu  Glu  Glu  Leu  Pro  Arg  Leu  Ala  Asp  Glu  Gly
20     25     30

Leu  Asn  Arg  Arg  Val  Ala  Glu  Asp  Leu  Asn  Leu  Gly  Asn  Leu  Asn  Val
35     40     45

Ser  Ile  Pro  Trp  Thr  His  Lys  Val  Gly  Asn  Phe  Thr  Gly  Leu  Tyr  Ser
50     55     60

Ser  Thr  Val  Pro  Cys  Phe  Asn  Pro  Lys  Trp  Gln  Thr  Pro  Ser  Phe  Pro
65     70     75     80

Asp  Ile  His  Leu  Gln  Glu  Asp  Ile  Leu  Asp  Arg  Cys  Lys  Gln  Phe  Val
85     90     95

Glu  Pro  Leu  Thr  Val  Asn  Glu  Asn  Arg  Arg  Leu  Lys  Leu  Ile  Met  Pr
100    105    110

Ala  Arg  Phe  Tyr  Pro  Asn  Val  Thr  Lys  Tyr  Leu  Pro  Leu  Asp  Lys  Gly
115    120    125

Ile  Lys  Pro  Tyr  Tyr  Pro  Glu  Tyr  Val  Val  Asn  His  Tyr  Phe  Gin  Thr
130    135    140

Arg  His  Tyr  Leu  His  Thr  Leu  Trp  Lys  Ala  Gly  Ile  Leu  Tyr  Lys  Arg
145    150    155    160

Gln  Ser  Thr  Arg  Ser  Ala  Ser  Phe  Cys  Gly  Ser  Pro  Tyr  Ser  Trp  Glu
165    170    175

Gln  Asp  Leu  Gln  His  Gly  Arg  Leu  Val  Phe  Gln  Thr  Pro  Lys  Arg  His
180    185    190

Gly  Asp  Lys  Ser  Phe  Cys  Pro  Glu  Ser  Pro  Gly  Ile  Leu  Pro  Arg  Ser
195    200    205

Ser  Val  Gly  Pro  Cys  Ile  Gln  Ser  Gln  Leu  Arg  Lys  Ser  Arg  Leu  Gly
210    215    220

Pro  Gln  Pro  Ala  Gln  Gly  Gln  Leu  Ala  Gly  Arg  Gln  Gln  Gly  Gly  Ser
225    230    235    240

```

Gly Ser Ile Arg Ala Arg Val His Pro Ser Ser Trp Gly Thr Val Gly  
 245 250 255  
 Val Glu Pro Ser Gly Ser Gly Pro Thr His Asn Cys Ala Ser Ser Ser  
 260 265 270  
 Ser Ser Cys Leu His Gln Ser Ala Val Arg Lys Ala Ala Tyr Ser Leu  
 275 280 285  
 Ile Ser Thr Ser Lys Gly His Ser Ser Ser Gly His Ala Val Glu Leu  
 290 295 300  
 His His Phe Pro Pro Asn Ser Ser Arg Ser Gln Ser Gln Gly Pro Val  
 305 310 315 320  
 Leu Ser Cys Trp Trp Leu Gln Phe Arg Asn Ser Glu Pro Cys Ser Glu  
 325 330 335  
 Tyr Cys Leu Cys His Ile Val Asn Leu Ile Glu Asp Trp Gly Pro Cys  
 340 345 350  
 Thr Glu His Gly Glu His Arg Ile Arg Thr Pro Arg Thr Pro Ala Arg  
 355 360 365  
 Val Thr Gly Gly Val Phe Leu Val Asp Lys Asn Pro His Asn Thr Ala  
 370 375 380  
 Glu Ser Arg Leu Val Val Asp Phe Ser Gln Phe Ser Arg Gly Asn Thr  
 385 390 395 400  
 Arg Val Ser Trp Pro Lys Phe Ala Val Pro Asn Leu Gln Ser Leu Thr  
 405 410 415  
 Asn Leu Leu Ser Ser Asn Leu Ser Trp Leu Ser Leu Asp Val Ser Ala  
 420 425 430  
 Ala Phe Tyr His Leu Pro Leu His Pro Ala Ala Met Pro His Leu Leu  
 435 440 445  
 Val Gly Ser Ser Gly Leu Ser Arg Tyr Val Ala Arg Leu Ser Ser Arg  
 450 455 460  
 Ser Arg Ile Asn Asn Asn Glu His Arg Thr Met Glu Asn Leu His Asn  
 465 470 475 480  
 Ser Cys Ser Arg Asn Leu Tyr Val Ser Leu Met Leu Leu Tyr Lys Thr  
 485 490 495  
 Tyr Gly Asn Lys Leu His Leu Tyr Ser His Pro Ile Ile Leu Gly Ile  
 500 505 510  
 Arg Lys Ile Pro Met Gly Val Gly Leu Ser Pro Phe Leu Leu Ala Gln  
 515 520 525  
 Phe Thr Ser Ala Ile Cys Ser Val Val Arg Arg Ala Phe Pro His Cys

530	535	540
Leu Ala Phe Ser Tyr Met	Asp Asp Val Val	Leu Gly Ala Lys Ser Val
545	550	555 560
Gln His Leu Glu Ser Leu Tyr Ala Ala Val	Thr Asn Phe Leu Leu Ser	
	565	570 575
Leu Gly Ile His Leu Asn Pro His Lys Thr Lys Arg Trp Gly Tyr Ser		
	580	585 590
Leu Asn Phe Met Gly Tyr Val Ile Gly Ser Trp Gly Thr Leu Pro Gln		
	595	600 605
Glu His Ile Val Gln Lys Ile Lys Met Cys Phe Arg Lys Leu Pro Val		
	610	615 620
Asn Arg Pro Ile Asp Trp Lys Val Cys Gln Arg Ile Val Gly Leu Leu		
	625	630 635 640
Gly Phe Ala Ala Pro Phe Thr Gln Cys Gly Tyr Pro Ala Leu Met Pro		
	645	650 655
Leu Tyr Ala Cys Ile Gln Ala Lys Gln Ala Phe Thr Phe Ser Gln Thr		
	660	665 670
Tyr Lys Thr Phe Leu Ser Lys Gln Tyr Leu Asn Leu Tyr Pro Val Ala		
	675	680 685
Arg Gln Arg Pro Gly Leu Cys Glu Val Phe Ala Asp Ala Thr Pro Thr		
	690	695 700
Gly Trp Gly Leu Ala Ile Gly His Gln Arg Met Arg Gly Thr Phe Val		
	705	710 715 720
Ser Pro Leu Pro Ile His Thr Ala Glu Leu Leu Ala Ala Cys Phe Ala		
	725	730 735
Arg Ser Arg Ser Gly Ala Lys Leu Ile Gly Thr Asp Asn Ser Val Val		
	740	745 750
Leu Ser Arg Lys Tyr Thr Ser Phe Pro Trp Leu Leu Gly Cys Ala Ala		
	755	760 765
Asn Trp Ile Leu Arg Gly Thr Ser Phe Val Tyr Val Pro Ser Ala Leu		
	770	775 780
Asn Pro Ala Asp Asp Pro Ser Arg Gly Arg Leu Gly Leu Tyr Arg Phe		
	785	790 795 800
Leu Leu Arg Leu Leu Tyr Arg Phe Thr Thr Gly Arg Thr Ser Leu Tyr		
	805	810 815
Ala Asp Ser Pro Ser Val Pro Ser His Leu Pro Asp Arg Val His Phe		

820

825

830

Ala Ser Pro Leu His Val Ala Trp Arg Pro Pro  
 835 840

## (2) INFORMATION FOR SEQ ID NO:3:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 400 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Met Gly Gly Trp Ser Ser Lys Pro Arg Lys Gly Met Gly Thr Asn Leu  
 1 5 10 15

Ser Val Pro Asn Pro Leu Gly Phe Phe Pro Asp His Gln Leu Asp Pro  
 20 25 30

Ala Phe Lys Ala Asn Ser Gln Asn Pro Asp Trp Asp Leu Asn Pro His  
 35 40 45

Lys Asp Asn Trp Pro Asp Ala Asn Lys Val Gly Val Gly Ala Phe Gly  
 50 55 60

Pro Gly Phe Thr Pro Pro His Gly Gly Leu Leu Gly Trp Ser Pro Gln  
 65 70 75 80

Ala Gln Gly Leu Leu Thr Thr Val Pro Ala Ala Pro Pro Pro Ala Ser  
 85 90 95

Thr Asn Arg Gln Ser Gly Arg Gln Pro Thr Pro Leu Ser Pro Pro Leu  
 100 105 110

Arg Asp Thr His Pro Gln Ala Met Gln Trp Asn Ser Thr Phe His  
 115 120 125

Gln Thr Leu Gln Asp Pro Arg Val Arg Ala Leu Tyr Phe Pro Ala Gly  
 130 135 140

Gly Ser Ser Ser Gly Thr Val Ser Pro Ala Gln Asn Thr Val Ser Ala  
 145 150 155 160

Ile Ser Ser Ile Leu Ser Lys Thr Gly Asp Pro Val Pro Asn Met Gln  
 165 170 175

Asn Ile Ala Ser Gly Leu Leu Gly Pro Leu Leu Val Leu Gln Ala Gly  
 180 185 190

Phe	Phe	Leu	Leu	Thr	Lys	Ile	Leu	Thr	Ile	Pro	Gln	Ser	Leu	Asp	Ser		
		195					200					205					
Trp	Trp	Thr	Ser	Leu	Asn	Phe	Leu	Gly	Gly	Pro	Thr	Val	Cys	Leu	Gly		
		210				215					220						
Gln	Asn	Ser	Gln	Ser	Gln	Ile	Ser	Ser	His	Ser	Pro	Thr	Cys	Cys	Pro		
225					230					235					240		
Pro	Ile	Cys	Pro	Gly	Tyr	Arg	Trp	Met	Cys	Leu	Arg	Arg	Phe	Ile	Ile		
				245					250					255			
Phe	Leu	Cys	Ile	Leu	Leu	Leu	Cys	Leu	Ile	Phe	Leu	Leu	Val	Leu	Leu		
			260					265					270				
Asp	Tyr	Gln	Gly	Met	Leu	Pro	Val	Cys	Pro	Leu	Ile	Pro	Gly	Ser	Thr		
		275					280						285				
Thr	Thr	Ser	Thr	Gly	Pro	Cys	Lys	Thr	Cys	Thr	Thr	Pro	Ala	Gln	Gly		
		290				295						300					
Thr	Ser	Met	Phe	Pro	Ser	Cys	Cys	Cys	Thr	Lys	Pro	Thr	Asp	Arg	Asn		
305					310					315					320		
Cys	Thr	Cys	Ile	Pro	Ile	Pro	Ser	Ser	Trp	Ala	Phe	Ala	Lys	Tyr	Leu		
				325					330					335			
Trp	Glu	Trp	Ala	Ser	Val	Arg	Phe	Ser	Trp	Leu	Ser	Leu	Leu	Val	Pro		
			340					345					350				
Phe	Val	Gln	Trp	Phe	Val	Gly	Leu	Ser	Pro	Thr	Val	Trp	Leu	Ser	Val		
		355				360						365					
Ile	Trp	Met	Met	Trp	Phe	Trp	Gly	Pro	Ser	Leu	Tyr	Asn	Ile	Leu	Ser		
		370				375					380						
Pro	Phe	Met	Pro	Leu	Leu	Pro	Ile	Phe	Phe	Cys	Leu	Trp	Val	Tyr	Ile		
385				390						395					400		

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 212 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(x) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met	Gln	Leu	Phe	His	Leu	Cys	Leu	Ile	Ile	Ser	Cys	Ser	Cys	Pro	Thr
1				5					10					15	



Leu Trp Gly Met Asp Ile  
30

Glu Leu Leu Ser Phe Leu  
45

Leu Leu Asp Thr Ala Ser  
60

Glu His Cys Ser Pro His  
75 80

Trp Gly Glu Leu Met Asn  
95

Asp Pro Ala Ser Arg Glu  
110

Gly Leu Lys Ile Arg Gln  
125

Phe Gly Arg Glu Thr Val  
140

Ile Arg Thr Pro Pro Ala  
155 160

Thr Leu Pro Glu Thr Thr  
175

Arg Arg Thr Pro Ser Pro  
190

Arg Arg Ser Gln Ser Arg  
205

Pro Leu Gly Ala Leu Pro Pro Ala Ser Pe  
35 40

His Gly Ala His Leu Ser Leu Arg Gly Le  
50 55

Ser Ala Gly Pro Cys Ala Leu Arg Phe Th  
65 70

Thr Thr Val Asn Ala His Gly Asn Leu Pe  
85 90

Thr Leu Gly Leu Ser Ala Met Ser Thr Th  
100 105

Lys Asp Cys Val Phe Asn Glu Trp Glu G  
115 120

Leu Lys Val Phe Val Leu Gly Gly Cys A  
130 135

Pro Ser Pro Cys Asn Phe Phe Thr Ser A  
145 150

(2) INFORMATION FOR SEQ ID NO:6

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 35 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6

ATAAGCTTATG CCCATATCTT ATCAACACTT CCGGA

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7

GAGTCTAGAC TGTTCCTAT TGTGA

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

Pro Ala Arg Asp Val Leu  
15

Gly Arg Pro Leu Pro Gly  
30

Pro	Leu	Gly	Ala	Leu	Pro	Pro	Ala	Ser	Pro	Pro	Val	Ile	Pro	Thr	Asp
		35					40					45			
His	Gly	Ala	His	Leu	Ser	Leu	Arg	Gly	Leu	Pro	Val	Cys	Ala	Phe	Ser
	50					55					60				
Ser	Ala	Gly	Pro	Cys	Ala	Leu	Arg	Phe	Thr	Ser	Ala	Arg	Arg	Met	Glu
	65				70					75					80
Thr	Thr	Val	Asn	Ala	His	Gly	Asn	Leu	Pro	Lys	Val	Leu	His	Lys	Arg
			85						90					95	
Thr	Leu	Gly	Leu	Ser	Ala	Met	Ser	Thr	Thr	Asp	Leu	Glu	Ala	Tyr	Phe
			100					105						110	
Lys	Asp	Cys	Val	Phe	Asn	Glu	Trp	Glu	Glu	Leu	Gly	Glu	Glu	Val	Arg
		115					120					125			
Leu	Lys	Val	Phe	Val	Leu	Gly	Gly	Cys	Arg	His	Lys	Leu	Val	Cys	Ser
	130					135					140				
Pro	Ser	Pro	Cys	Asn	Phe	Phe	Thr	Ser	Ala						
	145						150								

(1) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 35 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

ATAAGCTTATG CCCGTATCTT ATCAACACTT CCGGA

35

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 25 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

TAGTCTAGAC TCTGCTCTAT TGTGA

25

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 25 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GAGTCTAGAC TCGTGGTGA CTTCT

25

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 32 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

TGAGAATTCT CACGCTGGTC TCCATGCGAC GT

32

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

TTTGTTTACG TCCCT

10

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 27 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

CTAAGCTTAG TTCCGGAAG TGTGAT